2000

Determinants of Citations to Articles in Elite Reviews Interpreting Legal Citations

Ian Ayres

Fredrick E. Vars
University of Alabama - School of Law, fvars@law.ua.edu

DETERMINANTS OF CITATIONS TO ARTICLES IN ELITE LAW REVIEWS

*IAN AYRES and FREDRICK E. VARS**

ABSTRACT

This article analyzes the determinants of citations to pieces published from 1980 to 1995 in *Harvard Law Review, Stanford Law Review,* and *The Yale Law Journal.* We also rank articles by number of citations using regressions controlling for time since publication, journal, and subject area. To summarize a few of our results: citations per year peak at 4 years after publication, and an article receives half of its expected total lifetime citations after 4.6 years; appearing first in an issue is a significant advantage; international law articles receive fewer citations; jurisprudence articles are cited more often; articles by young, female, or minority authors are more heavily cited. Articles with shorter titles, fewer footnotes per page, and without equations have significantly more citations than other articles. Total citations generally increase with an article's length, but citations per published page peak at 53 pages.

I. INTRODUCTION

MANY legal citation studies have two shortcomings—they tend to focus too much on only the right-hand tail of the citation distribution and too little on the determinants of citation. The emphasis on the right-hand tail can be observed most clearly in Fred Shapiro's piece, "The Most-Cited Law Review Articles,"[1] which marked the path for later studies.[2] This right-tail

* Townsend Professor, Yale Law School, and J.D., Yale Law School, respectively. For helpful comments, the authors thank John Donahue, Ted Eisenberg, William Landes, Jim Lindgren, Peter Siegelman, David Steglich, Charles Vars, participants at the AALS law and economics section, and participants at the Interpreting Legal Citations symposium.

[1] See Fred R. Shapiro, The Most-Cited Law Review Articles, 73 Cal. L. Rev. 1540 (1985).

[2] See, for example, William M. Landes & Richard A. Posner, Heavily Cited Articles in Law, 71 Chi.-Kent L. Rev. 825 (1996); Deborah J. Merritt & Melanie Putnam, Judges and Scholars: Do Courts and Scholarly Journals Cite the Same Law Review Articles? 71 Chi.-Kent L. Rev. 871 (1996); Fred R. Shapiro, The Most-Cited Law Review Articles Revisited, 71 Chi.-Kent L. Rev. 751 (1996) (hereinafter Shapiro, Revisited); Fred R. Shapiro, The Most-Cited Articles from the Yale Law Journal, 100 Yale L. J. 1449 (1991). At least one other commentator has noted this bias. See Nancy Levit, Defining Cutting-Edge Scholarship: Feminism and Criteria of Rationality, 71 Chi.-Kent L. Rev. 947, 952 (1996) ("[P]erhaps the direction of quantitative inquiry is fundamentally wrong. We might learn more about scholarship and the politics of the legal academy if we explore instead what works are not being cited and why").

[*Journal of Legal Studies,* vol. XXIX (January 2000)]
 0047-2530/2000/2901-0020$01.50

bias is not unique to studies of law review articles. A recent study in this journal ranked the top 32 law schools on the basis of citations to articles written by professors at each of the law schools.[3] There have, however, been exceptions to the right-tail bias in legal citations analysis. A pair of articles examined citations to judicial opinions in order to identify the least significant United States Supreme Court justice.[4] And a sophisticated analysis of citations to opinions by judges on the federal courts of appeals included a complete ranking of the entire distribution of judges.[5] Notably missing from the literature is a comparable examination of the full distribution of law review articles (including the least cited).

Prior studies have also failed to delve very seriously into the determinants of citations. Studies of law review articles' citations have recognized the importance of opportunities to be cited but employ relatively crude statistical methods to correct for this effect. For example, in response to the age-of-article criticism of his all-time 100 most-cited articles' list, Shapiro added top-10 lists for articles published in each year from 1982 to 1991.[6] This methodology, however, does not allow for the simultaneous consideration of other variables.[7] A more sophisticated statistical approach is that of William Landes and Richard Posner, who recreated Shapiro's top-100 rankings, adjusting for the effects of article age and age squared.[8] Landes and Posner, however, did not correct for journal and subject area effects, which may influence opportunities to be cited. More fundamental, because they relied solely on cross-sectional data, Landes and Posner could not model the flow of citations over time or isolate the effects of changes in the

[3] See Theodore Eisenberg & Martin T. Wells, Ranking and Explaining the Scholarly Impact of Law Schools, 27 J. Legal Stud. 373 (1998). See also James Lindgren & Daniel Seltzer, The Most Prolific Law Professors and Faculties, 71 Chi.-Kent L. Rev. 781, 786 (1996) (ranking law school professors by the number of articles published in the 10 and 20 most-cited law reviews). Richard Posner has used citation analysis to examine the impact of Benjamin Cardozo and Learned Hand. See Richard A. Posner, Cardozo: A Study of Reputation (1990); Richard A. Posner, The Learned Hand Biography and the Question of Judicial Greatness, 104 Yale L. J. 511 (1994) (reviewing Gerald Gunther, Learned Hand: The Man and the Judge (1994)).

[4] See Frank H. Easterbrook, The Most Insignificant Justice: Further Evidence, 50 U. Chi. L. Rev. 481 (1983); David Currie, The Most Insignificant Justice: A Preliminary Inquiry, 50 U. Chi. L. Rev. 466 (1983). Easterbrook argues that Thomas Todd was less significant than Gabriel Duval[l].

[5] See William M. Landes, Lawrence Lessig, & Michael E. Solimine, Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges, 27 J. Legal Stud. 271 (1998). Consistent with the right-tail bias, however, much of the discussion centers on the most influential judges.

[6] See Shapiro, Revisited, *supra* note 2, at 773–77 (table 2).

[7] This criticism applies generally to Shapiro's comparisons between articles on his all-time and recent-year lists. See *id.* at 757–58.

[8] See Landes & Posner, *supra* note 2, at 830–31.

size of the data source.[9] And, to our knowledge, no prior studies have regressed article citations on a richer set of article and author attributes.

These two shortcomings are related because focusing on the "most cited" can obscure researchers' ability to identify the determinants of citations more generally. Steven Jay Gould, for example, pointed out that focusing solely on the decline of the .400 hitter in baseball led people to overlook explanations concerning the overall range of variation.[10] With regard to legal citations, it is at least plausible to argue that, among articles in elite law reviews, receiving few citations is a stronger indicator of relative quality than receiving many. Articles with many citations might be high quality or low quality (as other authors go out of their way to criticize the argument). Articles in elite law reviews with few citations, however, are more likely to be of low quality. Finding that a *Harvard Law Review* article has many fewer cites than other articles in the same subject area is hard to square with a hypothesis of excellence.[11]

The data in our study consist of 979 articles and other nonstudent pieces published in *Harvard Law Review, Stanford Law Review,* and *The Yale Law Journal* from 1980 to 1995.[12] We use these data to make three types of estimates:

Residual Rankings. Regressing article citations onto variables correlated with the opportunities the article had for citation (but independent of its quality), we rank the articles by how far their actual citations are above or below their expected citations (taking only opportunity variables into account). With some trepidation, we then list the most- and least-cited articles in our sample on the basis of these residual rankings.

Article and Author Attribute Regressions. By collecting panel data on the annual citations to particular articles, we are able to disentangle the expansion of our database, the Social Sciences Citation Index (SSCI), from the general tendencies of articles to be cited over time. We regress annual article citations on more than 60 regressors for article and author attributes.

Least-Cited Article Regressions. Returning to our residual rankings, we

[9] Some analyses of citations in areas other than law review articles employ more sophisticated statistical methodologies. For two of the best examples, see Landes, Lessig, & Solimine, *supra* note 5, and Eisenberg & Wells, *supra* note 3.

[10] See Steven Jay Gould, The Flamingo's Smile 218–21 (1985).

[11] Others have noted more general problems with inferring high quality from large citation counts. See, for example, Posner, Cardozo, *supra* note 3, at 70–73.

[12] We selected these three journals because these are generally considered to be the top three student-edited legal publications. Shapiro reports that these three journals published more of the most-cited articles from 1982 to 1991 than any other journal. See Shapiro, Revisited, *supra* note 2, at 763. The Appendix provides a more complete description of the SSCI database as well as the data we collected on article and author attributes.

explore which attributes were most associated with an article receiving few citations.

While this article marks a slight methodological advance over previous studies, there is no sense of triumph on our part. Identification problems plague all of our regressions. When we find that a particular attribute is positively correlated with citations to an article, should we conclude that (*a*) articles with this attribute have higher than average quality, (*b*) law review editors discriminate against articles with this attribute (demanding above average quality before publishing), or (*c*) subsequent authors discriminate in favor of citing this type of article (in the sense of citing to it more than its intrinsic merit deserves)? To this important question, our data speak not. And while we have criticized prior studies for not seriously investigating the determinants of citations, the difficulty of identifying the true meaning of many or few citations is so endemic that to our minds it calls into question the entire enterprise.

II. Opportunity Regressions and Residual Rankings

A raw citation count has several shortcomings as a measure of an article's impact or quality. The number of citations an article receives is obviously related to how many chances it has had to be cited. Thus, recent scholarship is penalized in compilations of most-cited articles.[13] Conversely, older pieces are less likely to have few citations because they have had so many more opportunities to be cited.

It is, however, possible at least conceptually to compare the relative quality of articles that have had different opportunities to be cited. We can regress total citations on variables that are correlated with the opportunities the article has had to be cited but are not correlated with its quality. The amount of time an article has been in print is the purest such variable, since time correlates positively with opportunity but is (much more) independent of an article's quality. The residuals of such an opportunity regression can then be used to rank the relative quality of different articles. These rankings crucially assume that, controlling for opportunities to be cited, better articles will be cited more often.[14] Articles with actual citations greater than predicted citations (that is, those with positive residuals) are likely to be

[13] This is not a new observation. For example, Landes and Posner used simple regression analysis to predict total expected citations of the articles on one most-cited list. See Landes & Posner, *supra* note 2. Our first regression mirrors theirs, except that we measure article age in months rather than years. See equation (1) *infra*.

[14] This of course will not always be true. See text around note 11 *supra* (discussing why few citations are stronger evidence of relative quality than many).

better than articles that lie below the opportunity regression line (that is, those with negative residuals).

To correct for the effect of article age, we estimated a loglinear regression predicting net total citations[15] based on age and age squared,[16] then ranked articles by standardized residuals.[17] The regression took the following form:[18]

$$\ln(\text{total citations}) = -10.459 + 5.245 \times \ln(\text{months}) - 0.497 \times [\ln(\text{months})]^2. \quad (1)$$

In this equation, the standardized residual of an article represents the difference (in standard deviations) between actual citations and the number of citations predicted for articles published in that month. In other words, the residual is the margin above or below the expectation for articles of the same age. Table 1 reports the five most-cited and five least-cited articles corrected in this way for time since publication.[19]

Table 1 underscores two possible additions to the opportunity regression. Criminal law and labor law each account for two of the bottom five articles listed. Conversely, the top of the list is heavy with discrimination law and legal theory articles. This suggests that, as we hypothesized above, subject area plays an important role in which articles get cited. Many more articles are written each year in constitutional law than in tax law, for example, so constitutional law pieces have many more opportunities to be cited. Thus, comparing constitutional law and taxation articles on the same scale is arguably unfair.[20] The second observation relates to the representation of the dif-

[15] We subtracted self-citations from gross total citations to obtain net total citations on the theory that self-citations do not reflect article quality. Hereinafter, we use "total citations" to refer to net total citations.

[16] Adding a "cubed age of article" variable to our regression models did not affect any of the results reported in Tables 1 and 2.

[17] Three articles received zero total citations. Because it is impossible to take a logarithm of zero, we added one divided by the sample size (1/530) to the total citations of every observation before converting to natural logs.

[18] All of the coefficients were significant at $p < 0.01$. The coefficients on ln(months) and $[\ln(\text{months})]^2$ imply that net citations increased quickest in early years then reached a maximum at 16.32 years.

[19] Several commentators have expressed horror about our publishing a list of least-cited articles. To demonstrate our good faith and to defray possible stigma, we note that a piece by one of the coauthors appeared near the bottom of the list: Ian Ayres & Joel Waldfogel, A Market Test for Race Discrimination in Bail Setting, 46 Stan. L. Rev. 987 (1994), appeared among the least-cited 10 percent of articles ranked by standardized residuals. Even worse, self-citations to this article actually exceeded citations by other authors (three versus two).

[20] The effects of size of a field on citations is perhaps ambiguous. More opportunities to be cited may also mean more competition for citations. See Landes, Lessig, & Solimine, *supra* note 5, at 323 (noting a comparable phenomenon in the context of judicial citations). In a particular subject area, articles in very elite law reviews have a competitive edge over other articles, so we think the opportunity effect will dominate the competition effect in our

TABLE 1

TOP FIVE AND BOTTOM FIVE ARTICLES RANKED BY STANDARDIZED RESIDUALS—AGE OF ARTICLE REGRESSION

Rank	Number of Citations	Reference
1	309	Charles R. Lawrence III, The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism, 39 Stan. L. Rev. 317 (1987)
2	297	Frank H. Easterbrook & Daniel R. Fischel, The Proper Role of a Target's Management in Responding to a Tender Offer, 94 Harv. L. Rev. 1161 (1981)
3	262	Robert M. Cover, The Supreme Court, 1982 Term—Foreword: Nomos and Narrative, 97 Harv. L. Rev. 4 (1983)
4	192	Angela P. Harris, Race and Essentialism in Feminist Legal Theory, 42 Stan. L. Rev. 581 (1990)
5	259	Mark V. Tushnet, Following the Rules Laid Down: A Critique of Interpretivism and Neutral Principles, 96 Harv. L. Rev. 781 (1983)
527 (tie)	1	George P. Fletcher, Mainfest Criminality, Criminal Intent, and the Metamorphosis of Lloyd Weinreb, 90 Yale L. J. 319 (1980)
527 (tie)	1	Lloyd L. Weinreb, Manifest Criminality, Criminal Intent, and the "Metamorphosis of Larceny," 90 Yale L. J. 294 (1980)
528	0	Zohar Goshen, Shareholder Dividend Options, 104 Yale L. J. 881 (1995)
529	0	Christopher T. Wonnell, The Contractual Disempowerment of Employees, 46 Stan. L. Rev. 87 (1993)
530	0	Benjamin Aaron, Unfair Labor Practices and the Right to Strike in the Public Sector: Has the National Labor Relations Act Been a Good Model? 38 Stan. L. Rev. 1097 (1986)

ferent journals on the list. *Harvard Law Review* includes three of the top five articles and none of the bottom five. *Harvard* has a larger readership than *Stanford* or *Yale*. For either prestige or distribution reasons, an article in *Harvard* may have received more citations than the same article would have received if it had appeared in one of the other two journals.

An alternative ranking would correct for subject area and journal effects. To this end, we generated journal dummy variables and a matrix of 13 subject area dummy variables.[21] Adding these variables to equation (1) signifi-

data. The observed regression coefficients are consistent with this hypothesis. See Table 3 *infra*.

[21] We generated two other opportunity variables: the first was equal to the sum of articles in each of the Wilson *Index to Legal Periodicals* subject areas (up to six) listed for each article, and the second was equal to the total number of articles in the SSCI database published in years after the article. The variables actually reduced the adjusted R^2 when added to equation (1), and neither obtained statistical significance, so we omit them from subse-

TABLE 2

TOP FIVE AND BOTTOM FIVE ARTICLES RANKED BY RESIDUALS—AGE OF ARTICLE, JOURNAL, AND SUBJECT REGRESSION

Rank	Number of Citations	Reference
1	240	Cass R. Sunstein, Interest Groups in American Public Law, 38 Stan. L. Rev. 29 (1985)
2	309	Charles R. Lawrence III, The Id, the Ego, and Equal Protection: Reckoning with Unconcious Racism, 39 Stan. L. Rev. 317 (1987)
3	80	Daniel A. Farber & Suzanna Sherry, Telling Stories out of School: An Essay on Legal Narratives, 45 Stan. L. Rev. 807 (1993)
4	158	Akhil Reed Amar, The Bill of Rights as a Constitution, 100 Yale L. J. 1131 (1991)
5	297	Frank H. Easterbrook & Daniel R. Fischel, The Proper Role of a Target's Management in Responding to a Tender Offer, 94 Harv. L. Rev. 1161 (1981)
526 (tie)	1	George P. Fletcher, Manifest Criminality, Criminal Intent, and the Metamorphosis of Lloyd Weinreb, 90 Yale L. J. 319 (1980)
526 (tie)	1	Lloyd L. Weinreb, Manifest Criminality, Criminal Intent, and the "Metamorphosis of Larceny," 90 Yale L. J. 294 (1980)
527	2	Michael C. Harper & Ira C. Lupu, Fair Representation as Equal Protection, 98 Harv. L. Rev. 1211 (1985)
528	0	Zohar Goshen, Shareholder Dividend Options, 104 Yale L. J. 881 (1995)
529	0	Christopher T. Wonnell, The Contractual Disempowerment of Employees, 46 Stan. L. Rev. 87 (1993)
530	0	Benjamin Aaron, Unfair Labor Practices and the Right to Strike in the Public Sector: Has the National Labor Relations Act Been a Good Model? 38 Stan. L. Rev. 1097 (1986)

cantly altered the results (see Table 2). For example, Cass Sunstein's article moves from off the most-cited list all the way to the top. The article benefited by appearing in *Stanford,* where expected citations (15.62) were significantly lower than in *Harvard* (26.58) or *Yale* (20.15)[22] and by being classified in the public law subject area, where articles had significantly fewer expected citations than the mean (17.87 versus 20.07).[23] At the other

quent analysis. We attribute the failure of these variables to the vagaries of Wilson classification and collinearity with age, respectively.

[22] $P = 0.000$ and $p = 0.033$, respectively. The reported values are derived from multiplying the regression coefficients by the mean values of the nonjournal variables, adding the appropriate journal coefficient, then converting from logarithms.

[23] $P = 0.017$. The other subject area dummies with statistically significant effects were criminal law (18.48 expected citations, $p = 0.072$), discrimination (36.57, $p = 0.054$), international law (14.86, $p = 0.016$), jurisprudence (45.36, $p = 0.000$), and feminism or Critical Legal Studies (CLS) (42.92, $p = 0.059$). (Recall that 20.07 citations is the mean.) A complete discussion of the determinants of citations is reserved for the next section. See Section III

end of the distribution, a *Harvard* article drops into the bottom five, attributable in part to the much higher expected citation count for articles in that law review.[24]

There is at least one very important criticism of our approach.[25] The journal and subject dummies may not be truly independent of article quality. To the extent the *Harvard* effect is attributable to higher average article quality rather than wider circulation or readership, it is problematic to include the journal dummies. To do so is to hold articles to a higher standard precisely because they were of high enough quality to appear in the top law review—an unfair ratcheting up. The same criticism could apply to the subject dummies. It may be that constitutional law articles are of higher average quality than international law articles. If this were true, allowing for higher expected citations for constitutional law articles would constitute an unfair penalty with respect to international law. Once we go beyond controlling for age,[26] it becomes difficult to identify variables that correlate with opportunity but that are independent of quality.

III. Determinants of Citations

The previous section reported citation rankings adjusted for the opportunities an article had to be cited. The opportunity variables included months

infra. Our ranking of the least-cited articles is not robust to alternative subject matter definitions and/or aggregations. For example, it may be that feminist history articles are cited substantially less than feminist theory articles. If this were true, the level of aggregation that we used would unfairly predict higher numbers of citations for all feminism pieces and therefore estimate a negative residual for feminist history articles.

[24] The article by Harper and Lupu was also penalized for being classified as a discrimination article, among which expected citations were higher than the average. The phrase "Equal Protection" in the title accounts for this classification, but the article is fundamentally about labor law. More broadly, we readily concede that article attributes distinct from quality may have pushed the other five articles to the bottom of Tables 1 and 2. Goshen was a lecturer at Hebrew University, so the lack of citations to his article could, for example, represent reader bias against foreigners. The Weinreb article is a 25-page response to a book by Fletcher, and Fletcher's article is a 30-page rebuttal. The Aaron article appeared in a labor law symposium and spans a mere 26 pages. Wonnell's article is the third in labor law among the bottom six. This overrepresentation suggests fewer opportunities for labor law articles to be cited—perhaps a separate labor law dummy should have been included in the model.

[25] There are at least two other ways to control for time since publication: (1) by grouping articles by the year published and (2) by limiting citation counts to a constant number of years since publication. We explored the second methodology but rejected it when some of the overall most-cited articles appeared on the list of articles cited least in the first 4 years after publication. For example, Robert Cover's 1983 Harvard foreword received 262 total citations by 1997 but none until 1989. This extreme case cautions against drawing strong conclusions of low impact for relatively young articles (like the Goshen and Wonnell articles in Tables 1 and 2). Receiving few citations early on may be evidence that an article was "ahead of its time."

[26] While it seems less plausible that the quality of articles has varied systematically over the time period studied, even the age and age-squared variables are at least potentially subject to this objection.

since publication, journal, and subject area. In this section, we explore the effects of other article and author characteristics on observed citations. Before reporting our results, a methodological point is in order. Citations in a given year is the outcome variable in this section, rather than total citations over the life span of the piece.[27] This obviously has the effect of greatly increasing the total number of observations. To exploit the panel structure of our data, we estimated a loglinear AR(1) panel model,[28] controlling for a variety of other article and author effects.[29] Estimating a panel model made it possible to distinguish between citing year effects and the true effects of article age. A final advantage of this approach is that we can model the flow of citations to a publication and test whether this pattern itself has shifted over time. To our knowledge, this approach is novel in the analysis of legal citations.

In fact, we find evidence of significant growth in the average number of citations over time. Figure 1 reports the estimated median citations for the years (from 1981 to 1997) in which the citations to a piece occurred.[30] Especially in light of the fact that the model controlled for age of article when cited, this result suggests that the SSCI database was larger in 1989 and subsequent years. Including citing year dummy variables in our model deflates the observed effects of other variables for expansion of the SSCI database.[31] This is a less constrained opportunity correction than including variables for age and age squared.

An obvious determinant of citations in a particular year is the number of years that have passed since the cited piece was published. After a lag period for the article to be read and for other citing pieces to be published,

[27] Unfortunately, we were not able to extract self-citations on a yearly basis from the SSCI database. As a second-best solution, we deflated citations in each year using a multiplier equal to (total gross citations − self-citations)/total gross citations.

[28] This is a variety of generalized least squares (GLS) models, which use both cross-section and time-series data. See William H. Greene, Econometric Analysis 485–88 (1990). We experimented with fixed effects, random effects, unstructured correlation, and other autoregressive panel models before settling on the AR(1) model. Because citations to an article in adjacent years seem more closely correlated than citations between year 1 and year 5, for example, an autoregressive model seemed the most appropriate, and the AR(1) model achieved greater significance than AR(2). The magnitude and direction of the effects noted in the text did not vary substantially between the random effects and autoregressive models. Because most of our independent variables (author and article attributes) are time invariant, it was impossible to estimate meaningfully a fixed effects regression. Our attempts to estimate an unstructured correlation model failed to achieve convergence.

[29] See Table 3 *infra*.

[30] The coefficients on the citation year dummies for every year after 1988 are positive and statistically significant.

[31] Robert C. Ellickson, Trends in Legal Scholarship, in this issue, at 517, similarly deflates observed citations to correct for changes in the size of the database. See also Eisenberg & Wells, *supra* note 3, at 385 (using a comparable deflator equal to the number of documents citing Laurence Tribe or Richard Posner).

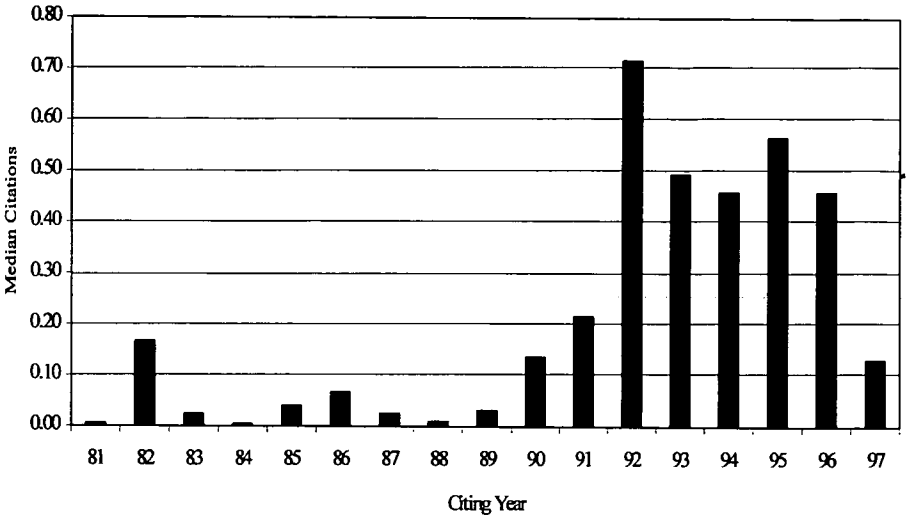

FIGURE 1.—Predicted median citations by citing year

we hypothesize that recent scholarship will be cited more than older work. Recent articles are more likely to be timely. Original ideas—which are presumably what these three elite journals publish—spread across legal scholarship, so that the original piece must compete for citations with more substitute pieces as time passes. If this hypothesis were true, we would expect to see citations per year rise quickly in the first few years after publication, crest, then ebb in subsequent years. Controlling for numerous article and author attributes, we observe exactly this pattern in median citations (see Figure 2).[32] Citations to a piece peaked 4 years after its publication, declined, then flattened out.[33] A simple tabulation reveals that half of total citations for all articles occurred before the articles were 4.61 years old.[34] This suggests that doubling citations about 4½ years after an article has

[32] Although the coefficients are not reported in Table 3, the regression model also included dummy variables for the number of years that elapsed between publication and an observed citation. From the coefficients on these variables we estimated the median citations depicted in Table 2.

[33] The flat distribution after year 7 is puzzling—one would expect citations to diminish to the right of the maximum. This pattern may be attributable to the growth of the SSCI database. See notes 29–30 and accompanying text *supra*. Because age of publication when cited and citing year are so highly correlated, including citing year dummy variables may not sufficiently correct for the expanded coverage of SSCI. Possible deficiencies aside, the observed flow of citations tells an interesting story about the impact of legal scholarship over time.

[34] An alternative methodology relies on the coefficients reported in equation (1). With a little algebra, these values imply a half-life of about 5 years.

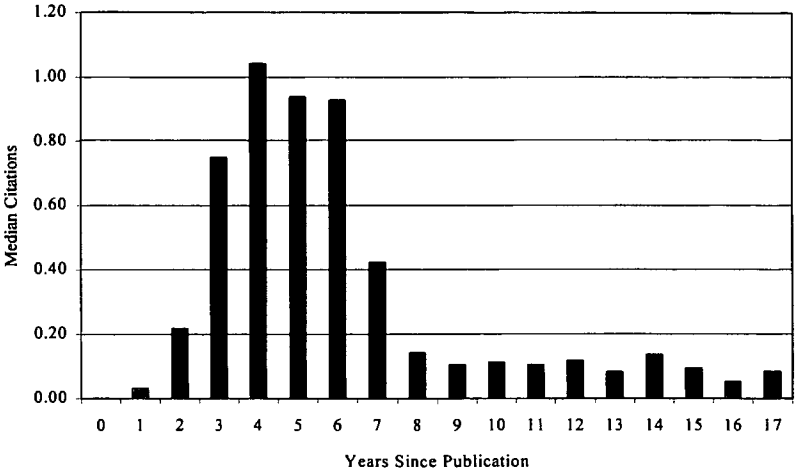

FIGURE 2.—Predicted median citations in a given year by age of article

been published would give an estimate of how many citations the article will receive over its life span.

In addition to time variables, we examine over 50 other article attributes. Many interesting findings emerge. *Harvard*'s advantage over *Stanford* and *Yale* disappears when citations are corrected for the other variables. The issue variables were not statistically significant, so the data do not support our hypothesis that articles appearing in issues filled during the traditional early spring and early fall submission waves would have been of higher quality (and garnered more citations) because they were drawn from a larger pool. The first piece in an issue received 108 percent more citations than pieces appearing fourth or later ($p = 0.013$).[35] To provide a sense of the substantive significance of this effect and the others reported below, the mean number of citations per year for the citation years included in Table 3 was 3.64 citations (the median value was one). Appearing second in issue was a marginally significant advantage ($p = 0.087$).[36]

Turning to subject effects, articles in the areas of jurisprudence and feminism or Critical Legal Studies (CLS) received significantly more than the average number of citations. The 78 percent bonus for articles in feminism and CLS was the largest positive effect. In contrast, international law and criminal law articles received significantly fewer citations than the aver-

[35] Coefficients in a loglinear regression represent the percentage effect of each variable on the average predicted outcome.

[36] See text around Table 4 *infra* for interpretation of the order-in-issue effect.

TABLE 3

RESULTS OF LOGLINEAR AR(1) PANEL MODEL PREDICTING CITATIONS IN A GIVEN YEAR

Variable	Coefficient	Z-statistic	Variable	Coefficient	Z-statistic
Article characteristics:			Table	.232	.803
Journal and issue:			Figure	.479	1.652
Stanford	−.035	−.146	Appendix	−.502	−1.759
Yale	.099	.495	ln(pages)	.869	.438
Issue 2	−.335	−1.130	ln(pages)^2	−.003	−.013
Issue 3	.089	.298	ln(footnotes per page)	−.482	−2.483
Issue 4	.264	.878	ln(footnotes per page)^2	−.150	−2.236
Issue 5	.117	.385	Colon in title	.131	.622
Issue 6	−.276	−.882	ln(characters in title)	−.120	−.621
Issue 7	.580	1.564	Author characteristics:		
Issue 8	.179	.435	Demographic:		
First in issue	1.079	2.477	White female	.467	2.074
Second in issue	.756	1.712	Minority male	.542	1.412
Third in issue	.407	.695	Minority female	1.640	2.546
First and nonalphabetic	.262	1.063	Aged 23–35	.833	3.547
Second and nonalphabetic	.326	1.170	Aged 36–40	.295	1.307
Third and nonalphabetic	−.229	−.316	Aged 47–82	.448	1.618
Subject:			Employment:		
Business	.174	.688	Chair	−.003	−.012
Common law	.236	.899	Professor	.448	1.612
Constitutional law	.120	.553	Associate professor	.328	1.025
Criminal law	−.675	−2.310	Assistant professor	.000	.001
Discrimination	.254	.905	Judge	.657	.527
Economics	.158	.627	Attorney	−.133	−.256
Feminism or Critical Legal Studies	.709	1.700	Local tenure piece	−.228	−.585
International law	−1.739	−4.623	At school of journal	.512	2.164
Jurisprudence	.781	3.517	Visited school of journal	−.139	−.403
Legal profession	.374	1.085	At other top-three school	.465	1.466
Procedure	−.236	−.992	At top-four to top-15 school	.671	3.110
Public law	−.320	−1.236	Prior experience:		
Miscellaneous	−.586	−1.054	Degree from school of journal	−.062	−.331
Other attributes:			Ph.D.	−.080	−.328
Foreword	.063	.112	On law review	.320	1.816
Coauthor	.581	2.142	U.S. Supreme Court clerk	.066	.316
More than two authors	−2.091	−2.242	Admitted to the bar	−.393	−1.534
Equation	−1.311	−3.353			

NOTE.—Observations = 5,431; N = 530; probability > χ^2 = 0.000.

age.[37] Among other article attributes, *Harvard* forewords were not significantly more cited than other articles. Coauthored articles were cited more frequently than single-author pieces, but pieces with more than two authors suffered a large penalty. Reporting an equation reduced average citations by an estimated 131 percent ($p = 0.001$). Articles with figures did 48 percent better than average, whereas articles with appendices did 50 percent worse than average; both effects were marginally significant. The significant negative coefficients on the footnotes-per-page variables suggest that articles with fewer footnotes were cited more frequently than articles with more footnotes, holding number of pages constant.

Characteristics of the author also played a significant role in determining citations. White female authors received 57 percent more citations than white men (the omitted category) ($p = 0.038$). Articles by minority women were the most heavily cited, with 164 percent more citations than articles by white men ($p = 0.011$). In a separate regression, we tested the hypothesis that citations to articles by minority authors increased after Richard Delgado's widely publicized article in 1984, which accused nonminority authors of discriminating against articles by minorities.[38] We constructed a dummy variable equal to one for pieces by minority scholars cited in years after 1984. There was no statistically significant evidence of increased citation to minority-authored articles after 1984.

Authors in the youngest quartile of the distribution (below 36 years of age) received significantly more citations than authors in the omitted quartile (ages 41–46). The first six variables characterizing the employment status of authors suggest the following ranking of authors from highest to lowest average citations: judges, full professors, associate professors,[39] assistant professors, and attorneys. However, the lack of statistical significance on these variables tells a more important story than the relative size of the coefficients. This is at least weak evidence that law review editors did not systematically favor work by full professors.

The next set of variables were designed to test for nepotism in the law review selection process. The "local tenure piece" variable is equal to one for associate professors publishing at the law journal where they teach. The observed negative coefficient on this variable supports the hypothesis that

[37] It should be noted that each piece was allowed more than one subject classification. For example, Ayres & Waldfogel, *supra* note 19, was classified in criminal law, discrimination, and economics.

[38] See Richard Delgado, The Imperial Scholar: Reflections on a Review of Civil Rights Literature, 132 U. Pa. L. Rev. 561 (1984).

[39] Harvard Law School does not have untenured associate professors, so we classified as "associate professors" Harvard assistant professors with more than 3 years teaching experience when their article was published. Our aim was to capture plausible tenure candidates.

tenure pieces are accepted by the home journal at a relatively low quality cutoff. However, the result is not statistically significant. The negative but insignificant coefficient on the "visited school of journal" variable tells a similar story with respect to favoritism toward visiting professors. The variable "at school of journal" is equal to one for pieces with at least one author who was a teacher at the law school where the piece was published. For example, an essay by a Stanford professor in *Stanford Law Review* is coded one. The next variable similarly captures law teachers who were visiting professors at the school where the article was published within 3 years prior to its publication. "At other top-three school" was designed to be compared with the "at school of journal" variable and is equal to one for professors at one of the other two schools who publish in the third school's journal. For example, an article by a Yale professor published in *Harvard* is one for this variable. Contrary to the nepotism hypothesis, there are no significant differences among these coefficients, which vary from 0.46 to 0.67. Articles authored by professors at top-15 law schools were cited significantly more than the articles by authors at lower ranked schools. Finally, of all of our "prior author experience" variables, only prior membership on law review was a marginally significant advantage.

Including Nonarticles. Next, we expanded the data set to include all nonstudent-written work, excluding book reviews, tributes, and other very short pieces. When reporting least-cited pieces, it seemed unfair to include essays, commentaries, symposium pieces, and the like. However, we believe this type of scholarship is important and can give additional insight into the determinants of citations.[40] When nonarticles are included, the number of observations increased to 9,529, reflecting 979 published pieces.

Many of the observed effects remain unchanged, but several interesting differences emerge. The *Stanford* and *Yale* journal dummy coefficients were negative and highly significant when nonarticles were considered. This suggests that *Harvard*'s prestige may derive not from better articles but, rather, from higher quality (or simply fewer) shorter pieces. In addition to the same subject area effects as noted above, articles in constitutional law and the legal profession received significantly more citations, whereas public law pieces received significantly fewer than the average. The pages coefficient is significantly positive and the pages-squared coefficient significantly negative, with magnitudes implying increasing citations throughout the relevant range of pages but declining citations per page after page 53. Articles with shorter titles received significantly more citations than articles with longer titles. The Delgado effect is in the predicted direction—

[40] Indeed, the mean citations to nonarticles was 16.42 citations. This is not a trivial figure, even though it is well below the 37.25 mean for articles. Moreover, 26 nonarticles earned more than 50 citations each.

that is, more citations to pieces by minorities after 1984—but falls short of statistical significance ($p = 0.137$). Finally, former United States Supreme Court clerks received significantly more citations than average, and authors admitted to the bar received significantly fewer.

Changes over Time. Using year-specific citations data allows us to test for changes over time in the determinants of citations. (Again, we limit the analysis to articles only in order to avoid complications like a higher proportion of nonarticles in more recent years.) To test the null hypothesis that the flow pattern was constant across time, we generated 168 dummy variables for the interactions between the age of the article when cited and the year of publication. Including these variables in a regression model allows a different intercept for each unique pair of publication year and citing year. For example, articles published in 1985 are allowed a unique predicted citation average for the year 1990. These interaction variables are jointly significant at $p = 0.000$ (chi-squared test), suggesting that the pattern observed in Figure 2 varied significantly across publication years.

Especially given this finding on citation flows, it makes sense to consider whether the effects of other variables have changed over time. With this goal in mind, we constructed a dummy variable equal to one for citing years after 1990 and interacted it with the variables in the regression underlying Table 3.[41] As above, this more constrained model reveals that the flow of citations after 1990 was significantly different than in earlier years. The null hypothesis that the age by post-1990 interaction variables are all equal to zero is rejected at $p = 0.000$ (chi-squared test). From the signs on the coefficients, it appears that citations peaked earlier in the more recent period.

The effects of several article and author attributes are significantly different in the period since 1990. None of the subject area interaction coefficents, however, were statistically significant.[42] *Harvard* forewords received

[41] We generated interaction terms for every variable with two exceptions: the citing year dummies and the age-of-article dummies greater than seven. (Because our data cover only citations through the first part of 1997, this last set of dummies is equal to zero for articles published after 1989.)

[42] There is obviously more than one way to analyze subject area trends. Robert Ellickson, for example, has examined changes over time in the percentage of articles that use particular words. See Ellickson, *supra* note 31, at 517. We constructed a set of five dummy variables for articles using the phrases "transactions costs" or "Coase," "critical race theory" or "critical legal studies," "feminism," "table 1," and "civic republicanism." (Because Westlaw had complete coverage for our three journals only since November 1982, we had to drop 1,211 observations.) In a separate post-1990 interaction model (omitting our original set of subject variables and the "table" variable), we observe a significant positive coefficient on the post-1990 interaction with the Coase variables, suggesting increased impact in later years for scholarship in areas associated with these buzzwords. Indeed, Coasian articles were significantly more cited than others in the later time period. Civic republican articles received substantially more citations in the later period than in the earlier period, but neither the change over time nor the difference from the median was statistically significant. Feminist articles received significantly higher average citations in both time periods. For CLS or Criti-

fewer citations than other articles in the earlier time period, but more citations in later years, and the increase was marginally significant ($p = 0.053$). Associate and assistant professors received fewer citations than other authors in the earlier period. However, both classes of authors did significantly better after 1990. Full professors also gained relative to other authors, suggesting a general rise in citations for articles by legal academics. The effect of the "local tenure piece" variable after 1990 was negative and statistically significant ($p = 0.047$), consistent with the nepotism hypothesis. In the earlier period, however, the coefficient was positive but just shy of marginal statistical significance ($p = 0.106$). Being at a top-15 law school had no significant effect on citations in the earlier time period, but these authors fared significantly better after 1990. This suggests that the payoff to prestige or the quality of elite faculty was higher in later years. Finally, the law review advantage noted above was significant only in the post-1990 period.

IV. Determinants of Least-Cited Articles

This article began with the premise that the least-cited articles in the Harvard, Stanford, and Yale law reviews are likely to be of poor quality. The previous section explored the effects of various article and author attributes in estimating citations per year for all articles. In this section, we turn our attention back to the least-cited articles and examine whether the same variables predict which articles were likely to fall into the bottom 10 percent of the citation distribution.

Rather than use raw citation counts, we corrected for months since publication, journal, and subject area effects using the opportunity regression underlying Table 2. We constructed a dichotomous outcome variable equal to one for articles that fell into the bottom 10 percent of articles ranked by standardized residuals. Next, we estimated a logistic regression model predicting this outcome variable based on the panoply of independent variables listed in Table 3. Unfortunately, we were forced to return to simple cross-sectional analysis since the outcome variable takes on the same value in every citing year period. This reduced sample size to 441 observations, which greatly weakened the statistical power of the model. The pseudo-R^2

---

cal Race Theory (CRT) articles, none of the observed coefficients were significant. (For this combined variable, the rise of CRT probably compensated for the widely recognized decline of CLS. See Ellickson, *supra*.) Finally, articles with "table 1" received significantly more citations than other articles but only in the post-1990 period. Compare William M. Landes & Richard A. Posner, The Influence of Economics on Law: A Quantitative Study, 36 J. Law & Econ. 385 (1993) (concluding that law and economics was growing through 1980s and that its growth exceeded that of political theory, CLS, and feminist legal theory).

of the regression was 0.2438, but only a handful of coefficients were statistically significant.

The *Stanford* and *Yale* coefficients were −0.947 and −0.968, respectively, and the latter was marginally significant ($p = 0.089$). These coefficients correspond to odds ratios of 0.388 for *Stanford* and 0.379 for *Yale*. Because we constructed the outcome variable based on residuals from a regression equation that allowed for journal intercept shifts, these coefficients can be best understood as measuring the variance (not mean) of citations. The negative coefficients suggest that articles in the Stanford and Yale law journals had lower variance in citations, such that fewer of them fell below the 10 percent threshold. None of the issue effects were statistically significant, but the first-in-issue coefficient of −1.855 came close at $p = 0.104$ (odds ratio = 0.156). Similarly, only the international law subject coefficient, at 1.272, bordered on statistical significance (odds ratio = 3.568; $p = 0.110$). As with journal effects, because we previously corrected for differences in mean citations across subject areas, this result is best understood as showing that international law articles have a higher than average variance in total citations. Having a colon in the title significantly reduced the probability of falling into the bottom 10 percent (odds ratio = 0.363; $p = 0.083$). Turning to author characteristics, articles by women were significantly less likely to be among the least cited (odds ratio = 0.253; $p = 0.032$). Last, former law review members produced articles that were about one-third as likely as others to earn citation residuals at the bottom of the distribution (odds ratio = 0.361; $p = 0.020$).

V. The Identification Problem

In the introduction we alluded to the fundamental problem that our results are subject to multiple causal interpretations. We proposed three alternative explanations for observing a positive relationship between an article or author attribute and citation counts: (1) quality—articles with that attribute were of higher than average quality; (2) editor bias—law review editors required a higher minimum quality to publish (that is, discriminated against) that type of article; and (3) citing bias—subsequent scholars discriminated in favor of citing articles with this attribute in ways unrelated to the article's intrinsic quality.[43] At various points in our previous discussion, we tried to foreground particular explanations. For example, with regard to our discussion of the "local tenure piece" variable, we tried to interpret the

[43] The least-cited regression results are subject to all of these interpretations plus the possibility that the results relate to variance in citation rates. For example, the finding that having a colon significantly reduced the likelihood of an article falling into the least-cited decile might indicate merely that such articles have a lower variance in citation.

negative coefficient as evidence of nepotism (a form of editor bias). Now we step back to consider whether—given the identification problem—it is possible to distinguish among alternative causal stories.

To see the problem more concretely, consider our finding in Table 3 that articles by white women received 47 percent more citations than articles by white men. This finding could have at least three different causes. First, editors may be biased against articles by women—setting a higher quality threshold for their acceptance. Second, editors may set the same quality threshold for both male- and female-authored articles but, conditional on being above this threshold, the articles by women were of higher average quality. Third, if we relax the assumption that the number of citations reflects the intrinsic quality of the article, then readers might cite female-authored articles more not because of their higher quality but for some other reason. For example, citing authors might, as a theoretical matter, discriminate against articles by men. The category of "citing bias" encompasses a wide variety of criteria that differ from an article's quality. To provide another illustration, international law articles may be cited substantially less frequently than other articles because there are fewer opportunities each year for these articles to be cited (at least in the SSCI database).

Some progress, however, can be made in distinguishing among causal stories. An analysis of order-in-issue effects helps to distinguish between citing bias and the other two effects.[44] By interacting the order variables with a dummy variable equal to one if the articles in the issue were arranged nonalphabetically, we can distinguish the effect of higher readership due simply to appearing first from the effect of higher quality as manifested by the editors' decision to eschew alphabetical order to give a higher quality article priority.[45] When we modified the panel model regression reported in Table 3 to include new interaction variables, we found the order effects shown in Table 4. The coefficients suggest that alphabetized articles appearing first received 108 percent more citations than articles placed fourth or later. Since alphabetized article order is not systematically related to article quality, this positive and statistically significant coefficient is suggestive of citing bias. The first article has a billboard effect that gives it greater opportunities to be cited. The data suggest that appearing first (or second) in the issue—quite apart from editor prioritization—boosted expected median

[44] Our analysis of issue order is inspired by Scott Smart & Joel Waldfogel, A Citation-Based Test for Discrimination at Economics and Finance Journals (Working Paper No. W5460, Nat'l Bureau Econ. Res. 1996).

[45] Of course, editors' quality judgments would be expected sometimes to coincide with straight alphabetical ordering. Without knowing more about the law review's ordering policies, however, we use nonalphabetical ordering as an admittedly underinclusive proxy for ordering based on perceived article quality.

TABLE 4

ORDER-IN-ISSUE EFFECTS—ARTICLES-ONLY REGRESSION

	Coefficient	*p*-Value
Variable:		
First in issue	1.079	.013
Second in issue	.756	.087
Third in issue	.407	.487
First and nonalphabetic	.262	.288
Second and nonalphabetic	.326	.242
Third and nonalphabetic	−.229	.718
	χ^2	*p*-Value
Difference:		
Alphabetical issues:		
First − second	2.250	.133
First − third	2.300	.129
Second − third	.610	.436
Nonalphabetical issues:		
First − second	.750	.386
First − third	4.570	.033
Second − third	3.020	.083

NOTE.—Observations = 5,431; N = 530.

citations. But the regression also suggests that nonalphabetized articles appearing first were likely to be cited even more—a full 134 percent more than articles appearing fourth or later. Within nonalphabetical issues, lead articles received significantly more citations than articles placed third (p = 0.033; chi-squared), which was not true for alphabetical issues. This suggests that, independent of any billboard effect, editors' choices of prioritization coincided with subsequent scholars' decisions of which articles to cite. This is at least consistent with the quality interpretation. Without being able to identify those issues in which alphabetical order happened to coincide with editors' quality judgments, it is impossible to estimate precisely the relative impact of each effect, but the data provide strong support for both interpretations.

Editor Bias. Editorial placement decisions may also be influenced by bias against certain categories of authors. For example, editors may systematically place inferior articles by men ahead of better articles by women. We also tested for bias against minorities, bias in favor of authors at the school of the journal, and bias in favor of authors at one of the top three law schools. To do so, we generated variables for the interactions between article placement and the characteristics of interest.[46] We know that first-

[46] To ensure that we were focusing on issues in which editors made discretionary placement decisions, we limited the data set to first and second articles in nonalphabetically ar-

placed articles generally receive more citations than articles appearing later. Thus, if female-authored articles placed second receive more citations than male-authored articles placed first, this would suggest that editors discriminated against women in their decisions on which articles to publish first in an issue. In fact, there was no evidence of either sexism or racism in placement decisions. The nepotism and elitism effects were actually the reverse of what we predicted: first-placed articles by outsiders and nonelite authors received fewer citations than second-placed articles by insiders and elite authors.[47] However, neither effect was statistically significant. This suggests that in placing articles law review editors may have favored authors who did not teach at the school of the journal or at the top three law schools generally.

Second, we tested the hypothesis that the top executive of a journal influenced acceptance decisions to favor articles by authors of his or her own racial or gender status.[48] To this end, we constructed three new interaction variables: female executive × female author, Asian executive × minority author, and black executive × minority author.[49] Negative coefficients on these variables would suggest a lower quality cutoff for articles written by authors with demographic characteristics matching the journal's executive. In fact, none of the interactions achieved statistical significance. The coefficient on the black interaction variable was, however, large and negative (−1.36), consistent with the favoritism hypothesis.

VI. Conclusion

There may be a strong temptation to read many of our results as recipes for citation success. Authors (or law review editors) might think that they could increase their citations if they just publish longer articles or shift toward publishing constitutional law pieces. Such inferences are fraught with peril. The fallacy of aggregation suggests that just because long articles have tended to be cited more in the past does not mean that journals should force authors to add 10 pages of pablum to their articles in order to generate

ranged issues with two or more articles. The result was 1,440 yearly observations for 121 articles.

[47] All of our findings here are consistent with those of Smart & Waldfogel, *supra* note 44, which examined the same types of bias in the placement of articles in economics journals.

[48] We use the generic term "executive" to refer to the editor-in-chief at Yale and the president at Harvard. The authors thank William Birdthistle, former managing editor of the *Harvard Law Review,* for gathering the data on presidents. Data for Stanford were unavailable.

[49] We subdivided executives into Asian and black on the basis of initial cross tabulations, which suggested differences between the two groups. Our sample included three black, four Asian, and eight female executives.

more citations. Still, with regard to an individual article, it is hard to resist thinking that affixing a shorter title might not increase the number of citations.[50]

Citations analysis also unavoidably has a gossipy and at times tawdry aspect. We are drawn to citation rankings—reading and discussing them around the water cooler—but we are simultaneously repulsed by them—criticizing their meanings (especially when we are excluded). Given the endemic identification problems discussed above, extreme modesty is in order. Indeed, while there is a certain ineffable value to these studies, we wonder whether this value justifies the time and journal pages devoted to the subject.[51]

APPENDIX

Description of the Data

Our source of citations data was the SSCI, compiled by the Institute for Scientific Information. We extracted citation, self-citation, and citation by citing year counts for every piece in the Harvard, Stanford, and Yale law reviews. The electronic version of the SSCI we used covered the period from 1980 to 1997, with only partial coverage in the first and last years. We used all of this citing information but limited the scope of cited pieces to the period from 1980 to the 1994–95 volume of each journal. Next, we examined in hard copy each piece on the list to determine how the journal classified it. We eliminated student-written work (notes, book notes, case notes, developments in the law, and so forth), tributes, correspondence, and book reviews. The remaining pieces included all articles, essays, commentaries, *Harvard* forewords, and symposium pieces. In order to a maintain a minimal degree of comparability among the pieces, we also eliminated all pieces with fewer than five pages and pieces shorter than 10 pages with no citations. In the end we were left with a total sample size of 979 pieces, including 530 articles.

From the hard copies of the pieces we created most of the article and author characteristic variables listed in Table 3. The professor, associate professor, assistant professor, judge, attorney, and age variables reflect the characteristics of the first author. For coauthored pieces, the other author indicator variables were also coded one if any of the coauthors satisfied the relevant criterion. The dagger notes were the primary source of these data. We also relied on the American Association of Law Schools (AALS) *Directory of Law Teachers* in each year from 1980 to 1995 to augment (and in some cases to create) the variables for visiting professor, local

[50] But if there were a general decline in the length of titles, even this effect might vanish. The general point is related to the Lucas critique of macroeconomic empiricism. Just because one observes that in the past the unemployment rate has declined when the money supply increases does not mean that the government can confidently manipulate the money supply in order to decrease unemployment. See Robert E. Lucas, Jr., Econometric Policy Evaluation: A Critique, in The Phillips Curve and Labor Markets 19 (Karl Brunner & Allan H. Meltzer eds., Carnegie-Rochester Conference Series on Public Policy No. 1, 1976).

[51] Compare J. M. Balkin & Sanford Levinson, How to Win Cites and Influence People, 71 Chi.-Kent L. Rev. 843 (1996).

TABLE A1

VARIABLE MEANS (and Standard Deviations for Continuous Variables) FOR ARTICLES

Variable	Mean	SD
Article characteristics:		
Citations:		
Total citations	38.8	47.0
Self-citations	1.5	2.2
Net citations	37.3	46.2
Journal and issue:		
Harvard	.294	
Stanford	.336	
Yale	.370	
Issue 1	.143	
Issue 2	.151	
Issue 3	.157	
Issue 4	.145	
Issue 5	.143	
Issue 6	.119	
Issue 7	.085	
Issue 8	.057	
First in issue	.566	
Second in issue	.332	
Third in issue	.064	
First and nonalphabetic	.113	
Second and nonalphabetic	.115	
Third and nonalphabetic	.025	
Subject:		
Business	.191	
Common law	.123	
Constitutional law	.272	
Criminal law	.077	
Discrimination	.121	
Economics	.145	
Feminism or Critical Legal Studies	.047	
International law	.047	
Jurisprudence	.228	
Legal profession	.058	
Procedure	.142	
Public law	.142	
Miscellaneous	.019	
"Transaction(s) costs" or "Coase"	.282	
"CLS" or Critical Race Theory ("CRT")	.022	
"Feminism"	.179	
"Table 1"	.101	
"Civic republicanism"	.055	
Other attributes:		
Harvard foreword	.028	
Coauthor	.132	
More than two authors	.006	
Equation	.064	

TABLE A1 (*Continued*)

Variable	Mean	SD
Table	.117	
Figure	.113	
Appendix	.098	
Pages	63.047	26.7
Footnotes	247.294	132.2
Colon in title	.391	
Characters in title	58.898	27.7
Author characteristics:		
Demographic:		
White male	.743	
White female	.179	
Minority male	.060	
Minority female	.017	
Age	40.158	8.428
Employment:		
Chair	.187	
Professor	.530	
Associate professor	.234	
Assistant professor	.100	
Judge	.004	
Attorney	.032	
Local tenure piece	.075	
At school of journal	.289	
Visited school of journal	.053	
At other top-three school	.074	
At top-four to top-15 school	.277	
Prior experience:		
Degree from school of journal	.311	
Ph.D.	.162	
On law review	.594	
U.S. Supreme Court clerk	.242	
Admitted to the bar	.847	

NOTE.—See note 42 *supra* for an explanation of the subject variables in quotation marks. $N = 530$.

tenure piece, Ph.D., law school attended, law review, United States Supreme Court clerk, and bar admission. The AALS *Directory* obviously covers only law teachers, and it is incomplete with respect to many of these variables, so we supplemented the data using the Martindale-Hubbell directories available in LEXIS and various directories of the judiciary. Nonetheless, many of these variables had to be coded as missing. We generated a set of dummy variables equal to one for missing observations, then recoded the missing values as zero. (We included this set of dummy variables in all analyses that included author characteristic dummies to retain these observations without biasing the results, but we do not report them because they are by definition impossible to interpret.)

The subject area variables were constructed differently. We linked our data to the Wilson *Index to Legal Periodicals* to obtain the Wilson subject codes (up to six) for each piece. These codes form the basis of our 13 subject area dummies.

Arbitrary grouping decisions were to a certain extent unavoidable. In addition, we searched (manually and electronically) titles and Wilson codes for key words associated with each of our subject areas. An alternative set of five subject area variables was calculated using a different methodology, which is described above.[52] For the 530 articles, Table A1 reports mean values for the variables of interest.

[52] See note 41 *supra*.